AD/A-001 102

A NEW APPROACH TO MULTI-STAGE STOCHASTIC LINEAR PROGRAMS

Richard C. Grinold

California University

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ORC 74-31	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER AD/A-001102
4. TITLE *(and Subtitle)* A NEW APPROACH TO MULTI-STAGE STOCHASTIC LINEAR PROGRAMS		5. TYPE OF REPORT & PERIOD COVERED Research Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Richard C. Grinold		8. CONTRACT OR GRANT NUMBER(s) N00014-69-A-0200-1055
9. PERFORMING ORGANIZATION NAME AND ADDRESS Operations Research Center University of California Berkeley, California 94720		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 047 120
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, Virginia 22217		12. REPORT DATE October 1974
		13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Stochastic Programming
Mathematical Programs
Infinite Horizon Optimization
Markov Chains
Conditional Expectations

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

(SEE ABSTRACT)

DD FORM 1473 1 JAN 73   EDITION OF 1 NOV 65 IS OBSOLETE

A NEW APPROACH TO MULTI-STAGE STOCHASTIC LINEAR PROGRAMS[†]

by

Richard C. Grinold
University of California, Berkeley

OCTOBER 1974 ORC 74-31

ABSTRACT

This paper considers an infinite stage linear decision
problem with random coefficients. We assume that the
randomness can be defined by a finite Markov chain. Under
certain assumptions we are able to calculate an upper bound
to an optimal value of the decision problem and to use that
bound to determine a useful initial decision.

A NEW APPROACH TO MULTI-STAGE STOCHASTIC LINEAR PROGRAMS

by

Richard C. Grinold

1. INTRODUCTION

This paper presents a novel and hopefully useful way of looking at multi-stage stochastic linear decision problems. We assume that the parameters that govern the evolution of the system are random variables with finite range and that the values of these parameters are determined by the state of a finite Markov chain. This assumption is a limitation on the general case of stochastic linear programs, however the loss in generality is offset by an ability to perform useful computations.

Our discrete time system can be viewed as a two stage decision process; the initial decision followed by all future decisions. The initial decision is subject to known constraints, leads to a known expected reward, and produces a random input into the second stage of the decision process. Our procedure, in effect, calculates an upper bound for the expected present value of the random input into the second stage of the decision process. If we make the assumption that this upper bound is a reasonable approximation for the value of the input, then we are able to calculate the initial decision that maximizes the expected first period reward plus the expected present value of all future decisions. Notice that the number of future decision stages is not important. In fact there can be an infinite number of future decisions.

Section 2 describes the decision process in detail while Section 3 defines the set of feasible policies. In Section 4 we see that each feasible policy leads to a sequence of conditional expectations of future decisions and a solution to an infinite horizon linear programming problem. In Section 5 we developed an

upper bound for the infinite horizon linear program and in Section 6 we use this knowledge to construct an initial feasible decision for the stochastic program. Section 7 is a summary and an outline of some interesting open questions.

The paper is based directly on two streams of thought. First, the study and solution of infinite horizon linear programs by Manne [6], Hopkins [5], Grinold and Hopkins [4], and Evers [1]. In particular, Assumption III in Section 5 and its consequences are based on ideas proposed by Evers [1]. The second idea is from earlier papers [2],[3] on dynamic stochastic decision processes in which the Markovian assumption was first proposed and exploited. An indirect and undoubtedly more important source has been the pleasure of learning from Roger Wets [7].

2. THE MODEL

This section presents a description of the stochastic decision process, introduces notation and defines terminology. The first part of the section describes a relatively simple model while the second portion of the section shows how more general models can be reduced to the same simple form.

We observe a system at discrete points in time $t = 0,1,2, \ldots$. At time t the system can be described by an $m + 1$ dimension vector (s_t, i_t) where $s_t \in R^m$ and $i_t \in \{1, 2, \ldots, k\}$. We refer to s_t as the *vector-state* and i_t as the *index-state*. The notation $\{s_t\}$ or $\{i_t\}$ refers to the sequence s_t or i_t for $t \geq 0$.

Given state (s_t, i_t) at time t the possible decisions u_t are constrained by

$$(1) \qquad A(i_t)u_t = s_t, u_t \geq 0$$

where $A(i)$ is an $m_i \times n$ matrix. Selection of a decision u_t results in a reward with expected present (time zero) value

$$(2) \qquad \beta^t c(i_t)u_t$$

where $\beta^t \geq 0$ is a discount factor.

The new index state at time $t + 1$ is determined by the transition probabilities of a finite Markov chain. Thus

$$(3) \qquad \text{Prob}[i_{t+1} = j \mid i_t = i] = P_{ij}.$$

Given the transition from i_t to i_{t+1}, the new vector state at time $t + 1$ is described by the linear relation

$$(4) \qquad s_{t+1} = K(i_t, i_{t+1})u_t.$$

To avoid a conceptual and theoretical difficulty we shall assume that the system cannot reach a situation where (1) has no feasible solutions.

Assumption I:

If $p_{ij} > 0$, and $v \geq 0$, then there exists a u satisfying

(5) $$A(j)u = K(i,j)v, \quad u \geq 0 .$$

We now consider several possible generalizations of the model described above and indicate how they can be reduced to the simple case.

First, suppose $i_t = i$, $i_{t+1} = j$ and $s_{t+1} = K(i,j)u_t + d(i,j)$. Then we could first expand the vector-state by one dimension and write

$$\tilde{s}_o = \begin{bmatrix} s_o \\ 1 \end{bmatrix}, \quad \begin{bmatrix} A(j) & 0 \\ \hline 0 & 1 \end{bmatrix} \begin{bmatrix} u_{t+1} \\ 1 \end{bmatrix} = \begin{bmatrix} K(i,j) & d(i,j) \\ \hline 0 & 1 \end{bmatrix} \begin{bmatrix} u_t \\ 1 \end{bmatrix} .$$

As a second variation, suppose

$$s_{t+1} = D(i,j)s_t + \tilde{K}(i,j)u_t .$$

Define $K(i,j) = D(i,j) A(i) + \tilde{K}(i,j)$; since $A(i)u_t = s_t$ we have

$$s_{t+1} = [D(i,j)A(i) + \tilde{K}(i,j)]u_t = K(i,j)u_t .$$

For a third extension, suppose $A(j)$ is $m_j \times n$ and

$$A(j)u_{t+1} = H(j)s_{t+1}$$

where $H(j)$ has full row rank. Then (by suitably interchanging columns) we can write $H(j) = B(j)[I,N(j)]$ where $B(j)$ has an inverse. Let $s = s_{t+1}$, and $u_{t+1} = u^1$ and partition $s = \begin{bmatrix} s^1 \\ s^2 \end{bmatrix}$, thus $H(j)s = B(j)s^1 + B(j)N(j)s^2$ so $B(j)^{-1}A(j)u^1 = s^1 + N(j)s^2$. Now expand the control space and define $u^2 - u^3 = s^2$. We have

$$\begin{bmatrix} B(j)^{-1}A(j) & -N(j) & N(j) \\ 0 & I & -I \end{bmatrix} \begin{bmatrix} u^1 \\ u^2 \\ u^3 \end{bmatrix} = \begin{bmatrix} s^1 \\ s^2 \end{bmatrix}$$

$$u^1, u^2, u^3 \geq 0 .$$

A fourth and trivial extension is to allow the dimension of u_t to vary with i_t, i.e., $u_t \in R^{n_i}$. However, we can compensate for this by setting $n = \max n_i$ and adding zero columns to the $A(i)$ and $K(i,j)$ matrices.

Finally, we can use all the Markov chain state space expansion tricks that transform apparently time dependent stochastic processes into Markov processes.

3. POLICIES

In this section we define the set of feasible policies and the value functions associated with them.

A realization of the system is defined to be the initial state (s_0, i_0) and the sequence i_t for $t \geq 1$. A policy is a function from the set of realizations to the set of decision sequences $\{u_t\}$. To be feasible a policy must have three properties.

(1)

   (i)   It is nonanticipative. Given two realizations with s_0 , and i_t for $0 \leq t \leq T$ identical, then the selection of u_T will be identical. In other words the value of u_T is independent of the values of i_t for $t > T$.

  (ii) The policy satisfies (2; 1).

  (iii) The policy satisfies (2; 4).

Note that the initial conditions (i_0, s_0) and the specification of a policy ψ are sufficient to recursively determine $\{u_t\}$ and $\{s_{t+1}\}$ for any realization $\{i_t\}$.

We shall let ψ denote a policy and Ψ the set of feasible policies. Given $s_0 = s$ and $i_0 = i$, we can define the value of a policy ψ .

(2)

   (i)
$$V^T(\psi, s, i) = E\left\{\sum_{t=0}^{T} \beta^t c(i_t) u_t\right\}$$

  (ii)
$$V^\infty(\psi, s, i) = \liminf_{T \to \infty} V^T(\psi, s, i) .$$

We can also define the optimal value functions.

(3)

   (i)
$$V^T(s, i) = \sup_{\psi \in \Psi} V^T(\psi, s, i)$$

  (ii)
$$V^\infty(s, i) = \sup_{\psi \in \Psi} V^\infty(\psi, s, i) .$$

4. CONDITIONAL EXPECTATIONS

Given a policy ψ and initial conditions s_o, and i_o we can define the expected decisions for all t conditional on the value of i_t. This section will show how these conditional expectations correspond to feasible solutions of an infinite horizon linear program.

Let

$$\pi_{ti} = \text{Prob}[i_t = i]$$

$$u_{it} = E[u_t \mid i_t = i]$$

(1)
$$s_{it} = E[s_t \mid i_t = i]$$

$$w_{it} = \pi_{ti} u_{it}$$

$$x_{it} = \pi_{ti} s_{it} \cdot$$

Note that

(2)
$$E[u_t] = \sum_{i=1}^{k} \pi_{ti} u_{it} = \sum_{i=1}^{k} w_{it} \cdot$$

From Bayes' rule we can calculate

(3)
$$\text{Prob}[i_t = i \mid i_{t+1} = j] = \frac{\pi_{ti} p_{ij}}{\pi_{t+1,j}}$$

where we interpret $0/0$ to be 0. Thus

(4)
$$s_{j,t+1} = \sum_{i=1}^{k} \frac{\pi_{ti} p_{ij}}{\pi_{t+1,j}} (K(i,j) u_{it}) = A(j) u_{j\,t+1}$$

or if we multiply (4) by $\pi_{t+1,j}$ we obtain

$$(5) \qquad x_{j,t+1} = \sum_{i=1}^{k} p_{ij} K(i,j) w_{it} = A(j) w_{j,t+1} .$$

Define

$$K = \begin{bmatrix} p_{11}K(1,1) & p_{21}K(2,1) & \cdots & p_{k1}K(k,1) \\ p_{12}K(1,2) & p_{22}K(2,2) & \cdots & p_{k2}K(k,2) \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ p_{1k}K(1,k) & p_{2k}K(2,k) & \cdots & p_{kk}K(k,k) \end{bmatrix}$$

$$x = \begin{bmatrix} x_{1t} \\ x_{2t} \\ \cdot \\ \cdot \\ \cdot \\ x_{kt} \end{bmatrix} \qquad w = \begin{bmatrix} w_{1t} \\ w_{2t} \\ \cdot \\ \cdot \\ \cdot \\ w_{kt} \end{bmatrix}$$

$$c = (c(1), c(2) \ldots, c(k))$$

and

$$A = \begin{bmatrix} A(1) & 0 & \cdots & 0 \\ 0 & A(2) & \cdots & 0 \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ 0 & 0 & \cdots & A(k) \end{bmatrix} .$$

With these definitions (5) becomes:

$$(7) \qquad x_{t+1} = K w_t = A w_{t+1} .$$

In addition, note that

$$V^T(\psi,s,i) = \sum_0^T \beta^t c w_t \ , \quad \text{since}$$

$$V^t(\psi,s,i) = E\left[\sum_0^T \beta^t c(i_t) u_t\right]$$

(8)

$$= \sum_0^T \beta^t \pi_{ti} \sum_{i=1}^k c(i) u_{it}$$

$$= \sum_0^T \beta^t c w_t \ .$$

We have demonstrated how each policy corresponds to a feasible solution of the horizon linear program;

(9)
$$\text{maximize} \quad \lim_{T \to \infty} \inf \sum_0^T \beta^t c w_t \ .$$

Subject to

$$A w_o = x_o$$

$$A w_t = K w_{t-1} \qquad t \geq 1$$

$$w_t \geq 0 \ . \quad t \geq 0 \ .$$

Although each policy defines a feasible solution to this infinite horizon linear program, the converse of this is not true. There exist feasible solutions of (9) that cannot be generated by a feasible policy $\psi \in \Psi$. Consider this example:

$$P = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$
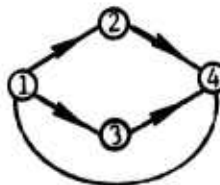
with

$$A(1) = A(2) = A(3) = K(1,2) = K(1,3) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

$$A(4) = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix}$$

$$K(2,4) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad K(3,4) = \begin{pmatrix} 2 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

and $K(4,1) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $s_o = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad i_o = 1$.

For this system the feasible solutions of the infinite linear program are

$t = 0 \qquad w_o = (w_{10}, 0, 0, 0)$ where $A(1)w_{10} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad w_{10} \geq 0$.

$t = 1 \qquad w_1 = (0, w_{21}, w_{31}, 0)$ where

$$A(2)w_{21} = \tfrac{1}{2}K(1,2)w_{10} = \tfrac{1}{2}A(1)w_{10} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad w_{21} \geq 0$$

$$A(3)w_{31} = \tfrac{1}{2}K(1,3)w_{10} = \tfrac{1}{2}A(1)w_{10} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad w_{31} \geq 0.$$

$t = 2 \qquad w_2 = (0, 0, 0, w_{42})^1$ where

$$A(4)w_{42} = K(2,4)w_{21} + K(3,4)w_{31} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad w_{42} \geq 0.$$

$t \geq 3 \qquad w_t = (0, 0, 0, 0)$.

Notice there is a feasible solution of the infinite horizon linear program with $w_{42} = (0,1,1)^1$.

Now let's examine the stochastic decision problem. At time 2, we shall surely be in index state 4, i.e., $\pi_{24} = 1$, and due to the special structure of the example the state vector can take one or two possible values, depending on the value of i_1 .

$$s_2 = \left(\begin{array}{l} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ with probability } \tfrac{1}{2} \\[2ex] \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ with probability } \tfrac{1}{2} \end{array} \right) .$$

A feasible policy must indicate what to do in either situation. Let $u_2(1)$ and $u_2(2)$ be the decision in either of the two cases. The expected decision is $u_{42} = \tfrac{1}{2}u_2(1) + \tfrac{1}{2}u_2(2)$. However, notice the first component of the three dimensional vector $u_2(1)$ must be positive. Thus if a feasible policy generates $w_{42} = \pi_{24}u_{42}$, the first component of w_{42} must be positive. We know, however, that a feasible solution of the infinite horizon linear program exists with the first component of w_{42} equal to zero. Thus all solutions of the infinite horizon linear program are *not* necessarily generated by feasible policies in the stochastic decision problem.

5. UPPER BOUNDS

In Section 4 we demonstrated that an infinite horizon linear program could be derived that must have an optimal value greater or equal to $V(s,i)$, the optimal value of the multi-state stochastic decision problem. Let $U(s,i)$ be the optimal value of the infinite horizon linear program as a function of the initial condition. We know that $V(s,i) \leq U(s,i)$. In this section we develop a linear program with optimal value $W(s,i)$ and show that under appropriate conditions $V(s,i) \leq U(s,i) \leq W(s,i)$.

Let Π be the set of all feasible solutions to the infinite horizon linear program, a sequence $\{w_t\} \in \Pi$ if and only if

$$Aw_o = x_o \ , \ w_o \geq 0$$

(1)

$$Aw_t = Kw_{t-1} \ , \ w_t \geq 0 \ , \ t \geq 1 \ .$$

Let $\tilde{\Pi} \subset \Pi$ be the set of solutions such that $\sum_{t=0}^{\infty} \beta^t w_t$ is finite. For $\{w_t\} \in \tilde{\Pi}$, define $w = \sum_{t=1}^{\infty} \beta^t w_t$ and note that

$$cw \leq U(s,i)$$

(2)

$$(A - \beta K)w = x_o$$

$$w \geq 0 \ .$$

Thus we can discover an upper bound on the value of the best solution in $\tilde{\Pi}$ by solving the finite linear program

$$\text{Max } cw$$

subject to

(3)

$$(A - \beta K)w = x_o$$

$$w \geq 0 \ .$$

Let $W(s,1)$ be the optimal value of (3). To show that $W(s,1)$ is an upper bound for $U(s,1)$ we must demonstrate that an optimal solution for the infinite horizon linear program can be found in the class $\tilde{\Pi}$. We must make two assumptions.

Assumption II:

An optimal solution to problem (3) exists.

Assumption III:

There exist (y,z) which satisfy

$$(4) \qquad y(A - \beta K) - z = c$$

$$z > 0 \text{ , and either } yA - c \geq 0 \text{ , or } yA \geq 0 \text{ .}$$

A consequence of Assumption III is that solutions not in $\tilde{\Pi}$ are infinitely bad. Let $\{w_t\} \in \Pi$, from (1) and (4) we can obtain for all T

$$
\begin{aligned}
yx_0 &= z \sum_0^T \beta^t w_t + \sum_0^T \beta^t cw_t + \beta^{T+1} yAw_{T+1} \\
&= z \sum_0^T \beta^t w_t + \sum_0^{T+1} \beta^t cw_t + \beta^{T+1}(yA - c)w_{T+1} \ .
\end{aligned}
$$

(5)

If Assumption III is satisfied, and $\{w_t\} \in \Pi \setminus \tilde{\Pi}$, then $z \sum_0^T \beta^t w_t \to +\infty$. However, nonnegativity of yA or $yA - c$ allows us to write

$$yx_0 \geq \sum_0^T \beta^t w_t + \sum_0^T \beta^t cw_t$$

(6)   or

$$yx_0 \geq z \sum_0^T \beta^t w_t + \sum_0^{T+1} \beta^t cw_t \ .$$

The term on the left is constant, the first term on the right diverges to $+\infty$, thus we must have the second term on the left diverging to $-\infty$. This indicates the value of any solution $\{w_t\} \in \Pi \setminus \widetilde{\Pi}$ is $-\infty$. Since, by Assumption II, the linear program has an optimal solution, we can conclude

(7)
$$V(s,i) \leqq U(s,i) \leqq W(s,i) .$$

6. OBTAINING DECISIONS FOR THE STOCHASTIC MULTI-STAGE LINEAR PROGRAM

This section indicates how the theory developed in Sections 2-5 can be used to generate a decision for the stochastic optimization problem. We know that $W(s,i)$, the optimal value of

$$\text{Max } cw$$

(1)
$$(A - \beta K)w = x_o$$

$$w \geq 0 ,$$

is an upper bound for $V(s,i)$ and we hope that $V(s,i)$ is close to $W(s,i)$. Even if this is so, an approximation of $V(s,i)$ is not extremely useful; we must determine an initial decision that is consistent and with that value. We show in this section how such an initial decision can be obtained.

Without loss of generality we can assume that $i_o = 1$, and that $p_{i1} = 0$ for $i = 1,2, ..., k$. Thus we cannot return to the initial index-state. In this case problem (1) becomes (for $k = 3$) :

Max.

$$c_1 w^1 + c_2 w^2 + cw^3$$

subject to

$$\begin{bmatrix} A(1) & 0 & 0 \\ -\beta p_{12}K(1,2) & A(2)-\beta p_{22}K(2,2) & -\beta p_{32}K(3,2) \\ -\beta p_{13}K(1,3) & -\beta p_{23}K(2,3) & A(3)-\beta p_{33}K(3,3) \end{bmatrix} \begin{bmatrix} w^1 \\ w^2 \\ w^3 \end{bmatrix} = \begin{bmatrix} s_o \\ 0 \\ 0 \end{bmatrix}$$

$$w^1 \geq 0 , \quad w^2 \geq 0 , \quad w^3 \geq 0$$

where $w^1 = \sum_{t=0}^{\infty} \beta^t w_{1t}$.

However, from the structure of K it is obvious that $w_{1t} = 0$ for $t \geq 1$, thus $w^1 = w_{10}$, a feasible initial decision that attains the upper bound $W(s,1)$.

7. SUMMARY AND OPEN QUESTIONS

This paper has presented a simple operational method of finding a good initial decision for a multi-stage stochastic programming problem and for calculating an upper bound on the optimal value of the stochastic program. The key modeling assumption is that the stochastic evolution of the system can be described by a Markov chain. The hope is that the upper bound is nearly exact and therefore the initial decision is nearly optimal. In a related paper [3], a special case was described in which the bound is exact, and the policy is optimal.

Several interesting related questions remain open. When does an optimal policy exist for the stochastic optimization problem? Is there an optimal stationary-Markov (u depends only on i and s) policy? When does the optimal solution of the infinite horizon linear program correspond to a realizable policy in the stochastic program?

REFERENCES

[1] Evers, J. J. M., "Linear Programming Over an Infinite Horizon,"
    University of Tilburg Press, Tilburg, The Netherlands, (1973).

[2] Grinold, R. C., "Manpower Planning with Uncertain Requirements,"
    ORC 73-29, Operations Research Center, University of California,
    Berkeley, (December 1973).

[3] Grinold, R. C., "A Class of Constrained Linear Control Problems with
    Stochastic Coefficients," CORE Discussion Paper 74-22.

[4] Grinold, R. C., and D. S. P. Hopkins, "Computing Optimal Solutions
    for Infinite-Horizon Mathematical Programs with a Transient
    State," Operations Research, Vol. 21, No. 1, (January-February 1973),
    p. 179.

[5] Hopkins, D. S. P., "Infinite Horizon Optimality in an Equipment Replacement
    and Capacity Expansion Model," Management Science, Vol. 18, No. 3,
    p. 145, (November 1973).

[6] Manne, A. S., "Sufficient Conditions for Optimality in an Infinite Horizon
    Development Plan," Econometrica, Vol. 38, p. 18-38, (1970).

[7] Wets, R., Lecture Notes - Stochastic Programming, University of California,
    Berkeley, (1967-68).